First Year Data Analysis

Prof. Dan Gillen
Yadong Lu
Student ID:75043108
Office: Bren Hall 2243

Contents

Analysis of Effect of Beta-carotene Supplementation on Beta-carotene Levels and Vitamin E Levels in the Serum

Yadong Lu [†]
Jun 30, 2017

Abstract

Beta-carotene supplementation is used as a preventative method for multiple cancers. This study analyzed the longitudinal data from 46 volunteers, aiming at investigating the pharmacokinetics of beta-carotene supplementation and how the supplementation dose affects other blood chemistries like vitamin E level. It is found that there is significant effect of beta-carotene supplementation on the time average serum beta-carotene levels (p value $< 10^{-3}$). And the effect is dependent on doses significantly(p value $< 10^{-3}$). Moreover, linear mixed effect model with random intercept is built and a likelihood ratio test was conducted. It is shown that there is significant effect of beta-carotene supplementation on both intercept and slope of serum beta-carotene levels (p value $< 10^{-3}$). However, there isn't enough evidence showing the effects of beta-carotene supplementation on the trajectory of the serum beta-carotene levels are different among doses (p value: 0.348). Similarly, effect of supplementation is found significant in serum vitamin E levels(p value $< 10^{-3}$). But the effect of supplementation does not significantly dependent on doses. A secondary analysis using baysian inference is performed. Overall, we conclude there is enough evidence suggesting that serum vitamin E levels are associated with serum beta-carotene levels over time.

Keywords: Beta-carotene supplementation, Serum vitamin E vevels, Linear mixed effect model, Baysian inference.

1 Introduction

Beta-carotene is a member of the carotenes and an antioxidant, which is commonly found in plants, whole grains and fruits such as carrots, spinach and sweet potatos. In the body, beta-carotene converts into vitamin A. Beta-carotene and other carotenoids provide approximately 50% of the vitamin A needed in the American diet.

Overwhelming observational evidence has existed that supports an association between lower lung cancer risk and greater consumption of carotenoid rich foods and, specifically, higher beta carotene intake [1],[2],[3]. And it is suggested that certain micronutrients might be inhibitors of cancer. Hence beta-carotene supplementation is commonly used as a preventative method for multiple cancers.

Therefore it would be our interest to first explore the pharmacokinetics of beta carotene supplementation before we start conducting large scale clinical experiments on investigating the role of beta carotene.

Of particular interest in the study of pharmacokinetics of beta-carotene supplementation is how different dose levels affect the accumalation of beta-carotene in the serum over time. Besides, since both beta-carotene and vitamin E are lipid soluble, it has been hypothesized that increased level of beta-carotene might have an impact on serum vitamin E levels. Hence we are also interested in investigating whether the beta-carotene supplementation has any effect on the vitamin E levels.

Previous literature has identified several confounding and precision effects in the association between serum beta-carotene levels and dose of beta-carotene supplementation. Firstly, there are a few influencial factors that are associated with serum beta-carotene levels. It has been documented by [4] that beta-carotene levels were slightly higher in younger volunteers, and were higher in female than in male. The study[4] was conducted on a total of 12,741 volunteers before any supplementation of beta-carotene. Additionally researchers have shown that serum beta-

[†]Student ID: 75043108

carotene levels were associated with life nutritional habits and lifestyle. For example, smoking generally has an decreasing effect on serum beta-carotene levels.

It was noted by [5] that lower non-HDL cholesterol was associated with lower serum concentrations of beta-carotene. Further, population with higher BMI (Body Mass Index) are found to have lower serum concentrations of beta-carotene levels [6].

2 Materials and Methods

2.1 Data Collection

The data was collected on 46 volunteers who were randomly assigned to five different doses of beta-carotene, i.e. 0, 15, 30, 45, or 60 mg/day in a double blind way. These volunteers were assigned placebo for months 0, 1, 2, and 3, then randomzied and treated at radomization dose (five different doses) for the rest of the months up to 15 months. The corresponding month of the measurements were recorded. Not all of the volunteers completed all the visits. And there are some volunteers who had more than 1 visits within the same month. Additionally, AUC (time average) of both beta-carotene and vitamin E levels are recorded as well. Discriptions of each variable in the dataset are provided in the Appendix.

2.2 Scientific Goals

This study mainly focus on the following four Goals:

1. Determine whether the supplementation of beta-carotene is associated with increase of time average of serum beta-carotene levels and whether the effect of supplementation dependent on dose.

2. Determine whether the supplementation of beta-carotene influence the trajectory of beta-caratene over time and whether this influence dependent on different doses.

3. Determine whether beta-carotene supplementation affect the serum vitamin E level over time and whether this effect dose dependent.

4. Quantify association between the serum vitamin E levels and the serum beta-caratene level.

2.3 Statistical Methods

Model 1:

$$\begin{aligned} cauc_i =& \beta_0 + \beta_1 \cdot Dose15_i + \beta_2 \cdot Dose30_i + \beta_3 \cdot Dose45_i + \beta_4 \cdot Dose60_i + \beta_5 \cdot I_{male_i} + \beta_6 \cdot bmi.ind \\ &+ \beta_7 \cdot vauc_i + \epsilon_i \end{aligned}$$

Model 2:

$$\begin{aligned} bcarot_{ij} =& \beta_0 + \beta_1 \cdot Dose15_i + \beta_2 \cdot Dose30_i + \beta_3 \cdot Dose45_i + \beta_4 \cdot Dose60_i + \beta_5 \cdot I_{male_i} + \beta_6 \cdot bmi.ind + \\ & \beta_7 \cdot month_{ij} + \beta_8 \cdot vite_{ij} + \beta_9 \cdot Dose15_i \cdot month_{ij} + \beta_{10} \cdot Dose30_i \cdot month_{ij} \\ & + \beta_{11} Dose45_i \cdot month_{ij} + \beta_{12} Dose60_i \cdot month_{ij} + b_{i0} + \epsilon_{ij} \end{aligned}$$

Model 3:

$$\begin{aligned} vite_{ij} =& \beta_0 + \beta_1 \cdot Dose15_i + \beta_2 \cdot Dose30_i + \beta_3 \cdot Dose45_i + \beta_4 \cdot Dose60_i + \beta_5 \cdot I_{male_i} + \beta_6 \cdot bmi.ind + \\ & \beta_7 \cdot month_{ij} + \beta_8 \cdot bcarot_{ij} + \beta_9 \cdot Dose15_i \cdot month_{ij} + \beta_{10} \cdot Dose30_i \cdot month_{ij} \\ & + \beta_{11} Dose45_i \cdot month_{ij} + \beta_{12} Dose60_i \cdot month_{ij} + b_{i0} + \epsilon_{ij} \end{aligned}$$

To address the first question, we use a linear model (Model 1) to explore the association between time average of serum beta-carotene levels and supplementation of beta-carotene. Here we choose to adjust for male, bmi and chol, which are identified as precision variables in the introductory part. Note that each individual only has one value representing the area under the curve (AUC) of beta-carotene averaging over time. Hence there won't be within subject variation. Therefore we choose a linear regression model for simplicity and we can use robust variation estimator to obtain a post-hoc fix of standard errors. Also, noticing the since we want to test whether there is any difference between the amount of doses per day, we treat every dose value as a factor (i.e. we create a variable, $Dose15_i$, which is an indicator of whether the i-th subject is assigned to dose of 15 mg/day. Similarly, there are three other indicators:$Dose30_i, Dose45_i, Dose60_i$). And the random term is assume to follow $\epsilon_i \sim N(0, \sigma^2)$. Here

bmi.ind is the indicator of bmi > 25 (classified as overweight).

In order to test whether beta-carotene supplementation has an effect on time average of beta-carotene levels, we construct an F statistics to test $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ against $H_1 : \beta_j \neq 0$ for some j in 1 to 4. Moreover, in order to test whether there is differential effects of beta-carotene supplementation on beta-carotene levels in serum, we can test the null hypothesis $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4$ against H_1: there exist i, j such that $\beta_i \neq \beta_j$. Further, we checked the pearson residual to determine whether the variance assumption is reasonable and a Q-Q plot was drawn to check the normality assumption.

As for the second scientific goal, a main effect linear mixed effect model (Model 2) was built to investigate the influence of beta-carotene supplementation on the trajectory of beta-carotene levels in serum. Notice that month 0, 1, 2, and 3, all patients are on placebo. And different doses of supplementation are assigned starting from month 4. Therefore we average accross month 0 to 3 and treat the serum beta-carotene levels in first 3 months as baseline values. The linear mixed effect model allows for individual variation in the intercept as well as slope of the change of serum beta-carotene levels. And one of the pros of likelihood based methods is that we can utilize likelihood ratio test to determine whether there is significant effect of beta-carotene supplementation on the trajectory. The model is fitted with maximum likelihood method because we will not be able to conduct likelihood ratio test if we fitted with REML. Since we want to know how beta-carotene supplementation affects on both intercept and slope of beta-caratene levels in serum, we add an interaction term between slope and dose in our model to determine whether different doses affect the rate of change of serum beta-carotene levels.

Again, we can test $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ against the alternative to see whether there is significant effect of beta-carotene supplementation on the intercept of serum beta-carotene levels. And we can test $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4$ against H_1: there exist i, j such that $\beta_i \neq \beta_j$ to see if there is any difference among different the doses.

For the third scientific goal, a linear mixed effect model was built to determine the ralationship of beta-carotene supplementation and serum vitamin E levels accross time. Similar to question 2, likelihood ratio tests were conducted to see if determine the effect of beta-carotene supplementation and whether it is dose dependent.

As for the last scientific question, we summarized the results from model 2 and model 3. In particuler, we look at the corresponding coeficients of serum beta-carotene levels and vitamin E levels and determine whether they are significant in respective models. In addition, we did Bayesian inference on a linear mixed effect model considering serum beta-carotene as response and determine how the 95% credible interval of the coefficient of serum vitamin E levels is different from 0.

3 Results

3.1 Exploratory Data Analysis

In this part we perform a prelimary discriptive data analysis to explore the dataset and provide evidence of how different dose level of beta-carotene supplementation affect the build up of serum beta-carotene over time and how does the does level might influence on vitamin E in the serum. To begin with, a histogram containing all serum beta-carotene level measurements was drawn to understand how it is distributed. As it can be observed from figure 1, they are approximately normally distributed, which indicates that a linear regression model would be sufficient to investigate the association of time average beta-carotene level and beta-carotene supplementation. Also, we notice from the boxplots that there female volunteers tend to have higher time average of beta-carotene levels and vitamin E levels in serum. This confirms with our analysis in the introductory part that it is appropriate to add male indicator as a precision variable.

Moreover, in order to explore the individual trajectory of serum beta-carotene, spagetti plots of groups assigned to different supplementation dose were made. Each line represents the change of serum beta-carotene level over time of a particuler individual. There are several interesting facts can be observed from the plots. The serum beta-carotene level seem to be stable within the first 3 month (measurement 0, 1, 2, and 3) for all of the subjects. Starting from the fourth month, subjects assigned with beta-carotene supplementation experienced a pretty sharp increase in their serum beta-carotene levels, while the beta-carotene levels of subjects assigned with 0

dose of supplementation stayed similar to previous months. Besides, the differential effect of different doses are less observable. This indicates that beta-carotene supplementation have substantial effects on serum beta-carotene levels. However, different supplementation doses (0, 15, 30, 45, or 60 mg/day) seem to have similar effect on the increase of beta-carotene level in serum.

Another interesting fact we noticed from the spagetti plots is that among subjects assigned with beta-carotene supplementation, their serum beta-carotene levels experienced a sharp decrease at around month 12, immediately after most of them reached their peak at month 11. The reason of the decrease is not clear. It might be due to the fact that execess serum beta-carotene is converted to vitamin A. This problem can be addressed in further research but currently it is not the main question of this study.

As for the association between serum vitamin E levels and beta-carotene supplementation, we can observe from the spagetti plots that the supplementation seem to have a negative effect on the level of serum vitamin E levels, which coincides with our hypothesis. Similarly, we can not differentiate the effects among different supplementation doses of beta-carotene.

In order to explore the evidence of potential random intercept and random slope in the model 2, we fit a linear model for each individual and look at how their confidence interval vary from each other. From the plot we can find that there is clear random intercept effect. However, the randomness in the slope parameter are not so evident. Therefore we choose to include only random intercept in the linear mixed effect model (model 2).

Characteristics	0 mg/day	15 mg/day	30 mg/day	45 mg/day	65 mg/day	Total
Gender						
Male	81(60%)	82(55.41%)	48(32.88%)	65(52%)	68(46.9%)	344(49.21%)
Female	54(40%)	66(44.59%)	98(67.12%)	60(48%)	77(53.1%)	355(50.79%)
Age	56.415 (4.001)	56.432 (4.37)	57.41 (3.98)	55.984 (3.051)	55.917 (4.49)	56.448 (4.06)
BMI	25.88 (3.41)	25.93 (3.44)	25.784 (2.52)	25.051 (3.04)	24.518 (2.04)	25.441 (2.98)
bcarot	293.22 (298.61)	811.12 (605.10)	939.64 (637.70)	1033.18 (697.43)	1117.63 (725.08)	838.55 (674.87)
NAs						6
vite	134.21 (21.92)	135.05 (21.29)	135.13 (22.91)	134.53 (17.63)	134.56 (21.60)	
NAs						6
chol	216.86(25.51)	224.03(28.92)	215.55(33.71)	211.68(31.66)	232.96(33.3)	220.52(31.63)
cauc	388.98(431.44)	1141.23(320.11)	1304.52(273.02)	1330.59(285.61)	1529.97(241.29)	1141.05(503.62)
NAs						4
vauc	8.02(1.11)	7.84(0.78)	8.24(1.06)	8.01(0.47)	8.33(0.7)	8.09(0.88)
NAs						4
bacrot by month						
Month 0-3	390.89(474.22)	1002.6(384.16)	1199.11(264.51)	1187.38(305.24)	1556.33(448.2)	1063.16(533.62)
Month 4	363(419.7)	1129.4(372.59)	1294.33(273.57)	1302.88(183.53)	1445.89(216.34)	1103.24(488.74)
Month 5	519.6(691.57)	1284.9(650.77)	1469(393.81)	1277.1(413.92)	1462.5(298.44)	1207.84(604.21)
Month 6	369.89(334.64)	1323.11(418.29)	1286(326.8)	1271.5(237.22)	1554.22(299.93)	1158.43(523.02)
Month 7	364.62(273.74)	1161.89(249.27)	1343.44(281.11)	1318.25(219.1)	1497.44(215.95)	1150.88(461.58)
Month 8	337.71(179.63)	1192.25(288.43)	1396.56(331.21)	1287.5(366.17)	1438.33(243.49)	1163.8(477.97)
Month 9	234.5(91.75)	1100.25(333.79)	1208.67(188.61)	1351.57(502.42)	1543.22(576.37)	1095.27(583.2)
Month 10	188.27(78.85)	1383.67(776.08)	1774.22(647.53)	1880.88(762.14)	2140.56(682.48)	1407.24(935.77)
Month 11	155.78(105.41)	1233.12(604.05)	1616.14(856.87)	2133.25(866.78)	1949.11(422.01)	1394.83(938.47)
Month 12	303.38(305.14)	500.62(308.59)	650.89(370.12)	996(460.13)	917.12(455.72)	656.49(439.91)
Month 13	179.14(91.07)	394.38(293.43)	510.22(271.41)	752(413.56)	648.12(443.94)	492.05(361.74)
Month 14	165.25(61.57)	292.38(164.9)	449.62(225.47)	532.83(287.11)	509.62(313.17)	407.97(257.54)
Month 15	165.25(61.57)	292.38(164.9)	449.62(225.47)	532.83(287.11)	509.62(313.17)	407.97(257.54)

Table 1: Summary statistics of 46 volunteers. For categorical variable: Count(%); for continous variabel: Mean(Std.)

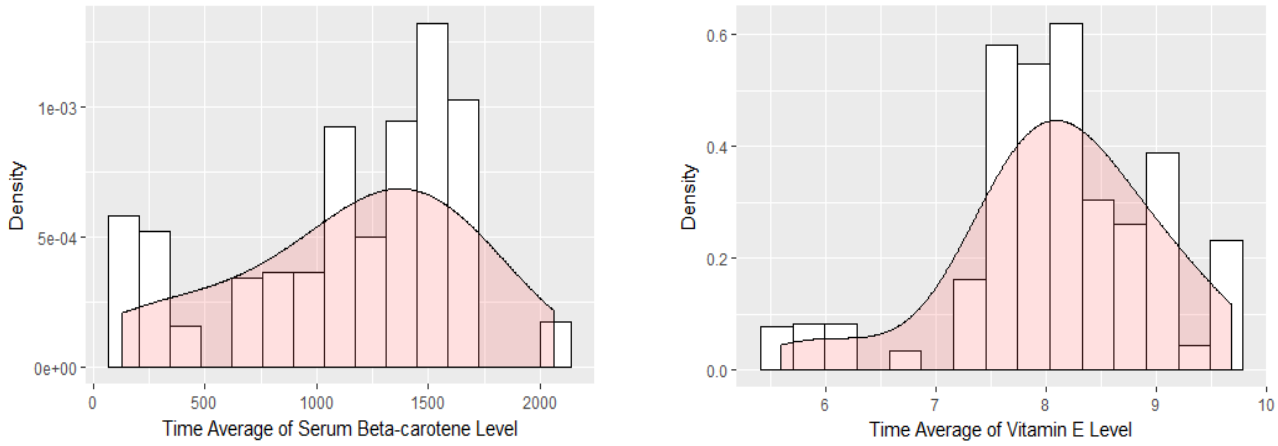

Figure 1: Histogram of time average serum beta-carotene levels and serum vitamin E levels

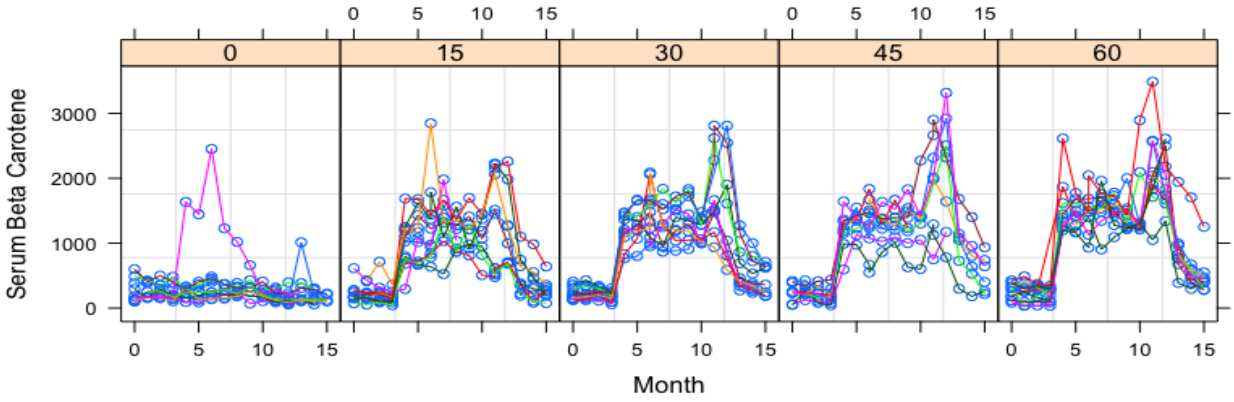

Figure 2: Spagetti plot of serum beta-carotene levels grouped by supplementation doses

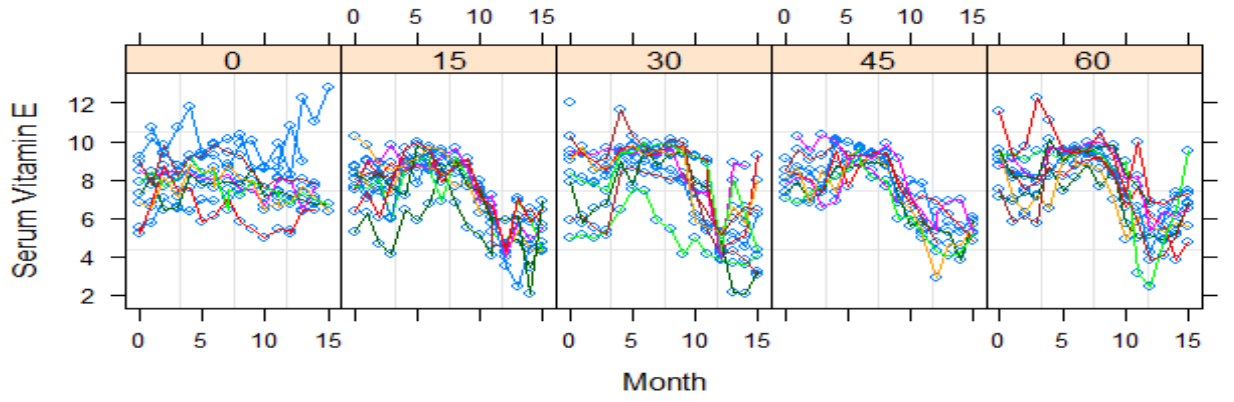

Figure 3: Spagetti plot of serum vitamin E levels grouped by supplementation doses

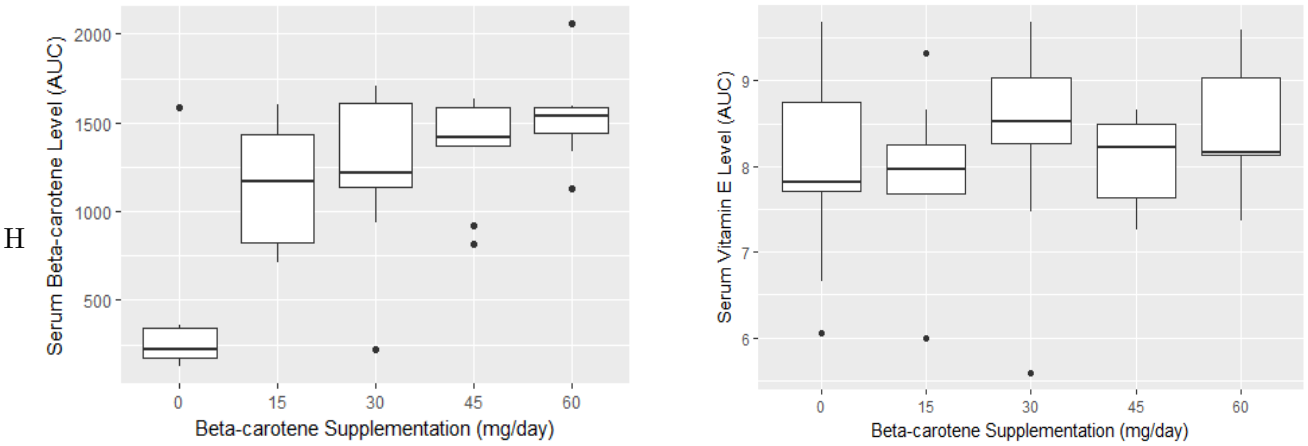

Figure 4: Boxplots of serum beta-carotene levels and serum vitamin E levels differed by doses.

In the data, we notice that serveral subjects have repeated measurements within one month. Some of the measurements within the same month are dubious. In particular, month 6, month 11 and month 12 have most of the repeated measurements. And for patient with ID number 1, his beta-carotene levels went from 1228 ug/mL to 2088 ug/mL within the month 6 and then dropped back to 1248 ug/mL at month 7. Such a variation is abnormal from other people's measurements. But we choose not to exclude this points from our dataset. Moreover, we found this measurement will not have significant effect in our inference as we shall see later.

Further, there is relatively small amount of missing data as we can see from table 1. And we would not worry too much the impact of missingness that could potentially bring to the inference. Therefore we choose to omit those data with missingness.

3.2 Inference

3.2.1 Linear Regression Model: Effects on Time Average of Beta-carotene Levels

In order to get consistent estimate, robust variance estimator is used to obtain confidence interval and p value.

	Est	robust ci95	Pr(>|t|)
(Intcpt)	275.332	(109.20, 441.46)	0.001
dose15	765.249	(694.84, 835.65)	0.000
dose30	849.577	(765.40, 933.75)	0.000
dose45	909.362	(840.28, 978.44)	0.000
dose60	1058.542	(977.24, 1139.84)	0.000
male	-208.070	(-251.93, -164.20)	0.000
bmi.ind	-181.402	(-222.16, -140.64)	0.000
vauc	41.805	(17.034, 66.58)	0.001

Table 2: Results of Model 1

As we can observe from the table, the estimated increase of the time average of serum beta-carotene levels (cauc) is 765.249 (95% robust CI: (694.84, 835.65), p value:$< 10^{-3}$) higher comparing subjects who were assigned to 15 mg/Day of beta-carotene supplementation to those who were assigned to 0 mg/Day controlling for gender, obesity status and time average of vitamin E level. And the increse of beta-carotene level seem to be bigger as the amount of dose increases.

Then we use ANOVA to conduct a test $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ against $H_1 : \beta_j \neq 0$ for some j in 1 to 4. The degree of freedom of the differences of residual sum of squares of reduced (under H_0) and full model is 4. The Resulting F statistic observed is $F_{obs} = 315.4$, with p value $< 2.2 * 10^{-16}$. Hence we can reject the null hypothesis that there is no effect of supplementation on the time average beta-carotene levels. To determine whether there is any difference between the doses, we conduct another test: $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4$ against H_1: there exist i, j such that $\beta_i \neq \beta_j$. Again, ANOVA is used and

the resulting F statistic observed is $F_{obs} = 34.793$, with p value $< 2.2 * 10^{-16}$. And the degree of freedom of the differences of residual sum of squares becomes 4.Hence we reject the null hypothesis that the effects of supplementation are equal. Hence we conclude that the difference of time average of serum beta-carotene level caused by difference supplementation doses is statistically significant. Also from the result, male, bmi status(bmi> 25) and vauc have significant effect (p value less than 0.05) in the response variable cauc.

Model diagnostics are performed and from the residual plot and qq plot (in Appendix), we observe that the residuals are approximately normally distributed around 0. No outlier is found in this process.

3.2.2 Linear Mixed Effect Model: Effects on Trajectory of Beta-carotene Levels

In this part, we built a linear mixed effect model (Model 2) with random intercept only in order to characterize the effect of beta-carotene supplementation on the trajectory of serum beta-carotene levels. The model fitting results are summarized as follows:

	Est	Std.	ci95.lo	ci95.hi	p value
(Intercept)	101.95	196.42	-283.03	486.93	0.60
month	-13.72	14.10	-41.35	13.91	0.33
vite	56.84	17.85	21.86	91.83	0.00
dose15	643.35	152.87	343.73	942.97	0.00
dose30	716.34	156.02	410.55	1022.13	0.00
dose45	631.67	161.46	315.21	948.13	0.00
dose60	828.19	156.21	522.02	1134.36	0.00
male	-182.30	72.91	-325.20	-39.40	0.02
bmi.ind	-133.54	73.34	-277.28	10.20	0.08
month:dose15	10.88	19.88	-28.08	49.84	0.58
month:dose30	13.46	19.82	-25.39	52.31	0.50
month:dose45	54.52	21.13	13.11	95.93	0.01
month:dose60	25.52	19.74	-13.17	64.21	0.20

Table 3: Results of Model 2

As we can observe from the table, supplementation of beta-carotene has significant effect on the intercept of the serum beta-carotene levels. For example, the expected difference of serum beta-carotene levels comparing patients assigned to dose 15 mg/day and dose 0 mg/day, controlling for other covariates is 643 ug/mL (95% CI: (343.73, 942.97), p value:$< 10^{-3}$). However, 10.88 (95% CI: (-28.08, 49.84), p value:0.58) is the expected difference in the slope comparing subjects assigned to dose 15 and subjects with dose 0, with similar level of other covariates, which is not considered significant.

Further, the likelihood result of ratio test ($H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_9 = \beta_{10} = \beta_{11} = \beta_{12} = 0$ against $H_1 : \beta_j \neq 0$ for some j) shows that the likelihood ratio statistic is 31.216 (degree of freedom: 8, p value:0.00013). Hence we reject the null hypothesis and conclude that there is significant effect of beta-carotene supplementation on the serum beta-carotene levels over time.

Moreover, we fit a reduced model assuming $\beta_1 = \beta_2 = \beta_3 = \beta_4$ and $\beta_9 = \beta_{10} = \beta_{11} = \beta_{12}$ (which is equivalent to replacing the 8 covariates related to dose by trt and trt:month model, where trt is an indicator of treatment, i.e. bacarot > 0) and compare its log likelihood with the full model. The resulting likelihood ratio statistic is 6.71 (degree of freedom: 6, p value: 0.348), which is considered to be not significant. Therefore we conclude that there isn't enough evidence showing the effects of beta-carotene supplementation on the trajectory of the serum beta-carotene levels are different among different doses.

3.2.3 Linear Mixed Model: Effects on Serum Vitamin E Levels over Time

In this section, a linear mixed effect model (Model 3) with random intercept only was built aiming at investigating how beta-carotene supplementation affects the serum vitamin E levels over time. The

result of model fitting is summarized in the table. As we can see, the serum vitamin E levels are decreasing with month. In fact, the serum vitamin E level decreases 0.07 ug/mL (95% CI: (0.01, 0.13), p value:0.03) by one month, controlling for other parameters in the model. And we notice that rate of decreasing is significantly larger when subjects are assigned to beta-carotene supplementation. For example, -0.27 (95% CI: (-0.35 , -0.19), p value: $< 10^{-3}$) is the expected change in the slope of vitamin E levels comparing subjects having 15 mg/day supplementation with subjects who don't have beta-carotene supplementation controlling for other covariates in the model. However, the change of rate are similar among different doses of supplementation.

Similar to the previous part. Testing whether there is any effect of beta-carotene supplementation is equivalent to test $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_9 = \beta_{10} = \beta_{11} = \beta_{12} = 0$ against there exsits some $\beta_j \neq 0$. The resulting likelihood ratio statistic is 34.691 (degree of freedom: 8, p value:3.045e-05). Hence we conclude that there is significant evidence showing beta-carotene supplementation is associated with serum vitamin E levels.

Additionally, a reduced linear mixed effect model assuming $\beta_1 = \beta_2 = \beta_3 = \beta_4$ and $\beta_9 = \beta_{10} = \beta_{11} = \beta_{12}$ is fitted. Then the likelihood ratio statistic is 2.276 (degree of freedom: 6, p value 0.892). Therefore we do not reject the null hypothesis. This indicates that there isn't enough evidence showing that the effects of beta-carotene supplementation is different depdending on doses.

	Est	Std.	95% CI	p value
(Intercept)	8.51	0.38	(7.76, 9.26)	0.00
month	-0.07	0.03	(-0.13, -0.01)	0.03
I(bcarot/1000)	0.32	0.10	(0.12, 0.52)	0.00
dose15	0.64	0.46	(-0.26, 1.54)	0.17
dose30	0.98	0.47	(0.06, 1.90)	0.04
dose45	1.01	0.49	(0.05, 1.97)	0.04
dose60	0.89	0.48	(-0.05, 1.83)	0.07
male	-0.09	0.26	(-0.60, 0.42)	0.73
bmi.ind	-0.43	0.26	(-0.94 ,0.08)	0.11
month:15	-0.27	0.04	(-0.35 , -0.19)	0.00
month:dose30	-0.29	0.04	(-0.37, -0.21)	0.00
month:dose45	-0.33	0.05	(-0.43, -0.23)	0.00
month:dose60	-0.25	0.04	(-0.33, -0.17)	0.00

Table 4: Summary of results from model 3

3.2.4 Association of Serum Vitamin E Levels and Serum Beta-carotene Levels and Secondary Anlysis

In model 2, the vitamin E level $vite_{ij}$ is added in the model. And from the result we know that 56.84 (95% CI: (21.86, 91.83), p value: 0.0015) is expected difference of serum beta-carotene level comparing subjects differing in $vite_{ij}$ of 1 ug/mL controlling for other covariates across time. Further in model 3, we found that the 0.32 (95% CI: (0.12, 0.52), p value: 0.0012) is the expeted difference in serum vitamin E level comparing subjects deffering in beta-carotene level of 1000 ug/mL with similar value of other covariates in the model. Therefore we would argue that there is relatively strong evidence showing that there might be association between serum beta-carotene levels and serum vitamin E levels over time.

Additionally, we conduct a secondary analysis we built a main effect linear mixed effect model with response variable to be serum beta-carotene levels. In this model, we add month, vite, ind.car (indicator of supplementation) and male, bmi and month:ind.car interaction. We fit this model using MCMC algorithm. The result shows that the 95% credible interval for the coefficient of serum vitamin E levels is (61.5617, 106.2152) which is significantly different from 0. The detailed result is attached in the appendix. Therefore we can conclude that there is enough evidence suggesting that there is significant association between serum vitamin E levels and

serum beta-carotene levels over time.

4 Discussion

In our study, we analyze the longitudinal data on serum beta-carotene level and serum vitamin E level of 46 volunteers to explore the pharmacokinetics beta-carotene supplementation and the its influence on serum vitamin E levels. We first explore the association between the beta-carotene supplementation and time average of serum beta-carotene levels. And it is found that the expected time average of beta-carotene is significantly higher in subjects with beta-carotene supplementation. For example, subjects assigned with 60 mg/day supplementation is 1058.542(95% CI: (977.24, 1139.84), p value: $< 10^{-3}$) higher in time average of beta-carotene level than subjects without supplementation, with same gender, similar BMI and time average of serum vitamin E levels. And the test results confirm that there is significant relation between supplementation and increase of time averaged serum beta-carotene level.

We quantify the influence of beta-carotene supplementation on trajectory of beta-carotene level using a linear mixed effect model (model 2). And it is found that patients assigned with 60 mg/day supplementation have the highest beta-carotene levels (estimate: 828.19, 95% CI: (522.02, 1134.36), p value: $< 10^{-3}$) in serum across time when controlling for other parameters. And subjects assigned with 60 mg/day supplementation have the highest rate of change (slope) in beta-carotene levels (estimate:54.52, 95% CI: (13.11, 95.93), p value: 0.01). From the result of the likelihood ratio test, we found that there is siginificant (p value: $< 10^{-3}$) effect of beta-carotene supplementation on the serum beta-carotene levels over time. However, there isnt enough evidence showing the effects of beta-carotene supplementation on the trajectory of the serum beta-carotene levels are different among doses(p value: 0.348).

Then a linear mixed effect model with intercept only (model 2) is fitted to explore the relation between beta-carotene supplementation and serum vitamin E levels over time. It is found that the serum vitamin E levels are decreasing with month. And beta-carotene supplementation correspond to higher rate of decrease in serum vitamin E levels. For example, For example, -0.27 (95% CI: (-0.35 , -0.19), p value: $< 10^{-3}$) is the expected change in the slope of vitamin E levels comparing subjects having 15 mg/day supplementation with subjects who don't have beta-carotene supplementation controlling for other covariates in the model. The test result show that there is siginificant (p value: $< 10^{-3}$) effect of beta-carotene supplementation on the serum vitamin E levels over time (p value:3.045e-05). But there isnt enough evidence showing that the effects of beta-carotene supplementation is different depdending on doses (p value 0.892).

From previous results of model 2 and model 3, we have seen that serum vitamin E levels and serum beta-carotene levels are associated with each other. In model 2, 56.84 (95% CI: (21.86, 91.83), p value: 0.0015) is expected difference of serum beta-carotene level comparing subjects differing in $vite_{ij}$ of 1 ug/mL controlling for other covariates across time. And in model 3, 0.32 (95% CI: (0.12, 0.52), p value: 0.0012) is the expeted difference in serum vitamin E level comparing subjects deffering in beta-carotene level of 1000 ug/mL with similar value of other covariates in the model. Therefore we conclude that serum vitamin E levels and serum beta-carotene levels are associated with each other over time.

This study has several limitations. Firstly, we only have 46 volunteers participating in this study. But in model 1, we used robust variance estimator as a post hoc fix of our stand errors, which typically requires more than 50 subjects to have stable performance. Further work could be conducted if we have more volunteers included in the study. In addition, more work could be done to select the prior distribution in the secondary analysis.

5 Reference

[1] $\beta-$ Carotene and lung cancer: a case study. Demetrius Albanes. *Am J Clin Nutr June 1999 vol. 69 No. 6 1345s-1350s*

[2] Anti-tumor activity of β-carotene, canthaxanthine, and phytoene. Mathews-Roth MM. *Oncology 1982 ;39 :33 7.*

[3] Epidemiologic evidence for -carotene and cancer prevention. Van Poppel G, Goldbohm RA. *Am J Clin Nutr 1995 ;62(suppl) :1393S 402S.*

[4] Factors influencing blood concentration of retinol, alpha-tocopherol, vitamin C, and bold italic beta-carotene in the French participants of the

SU.VI.MAX trial. S Hercberg and A Favier. *European Journal of Clinical Nutrition (2006) 60, 706717.*

[5] Human serum carotenoid concentrations are related to physiologic and lifestyle factors. Brady WE, Mares-Perlman JA, Bowen P, Stacewicz-Sapuntzakis M. *Journal of Nutr. 1996 Jan;126(1):129-37.*

[6] Associations Between Body Mass Index and the Prevalence of Low Micronutrient Levels Among US Adults. Joel E. Kimmons et. al.

6 Appendix

6.1 Diagnostic Plots

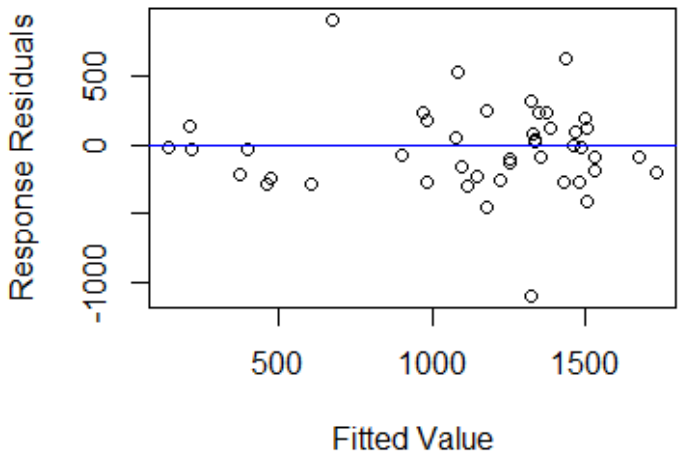

Figure 5: Response pearson residuals against the fitted value: model 1

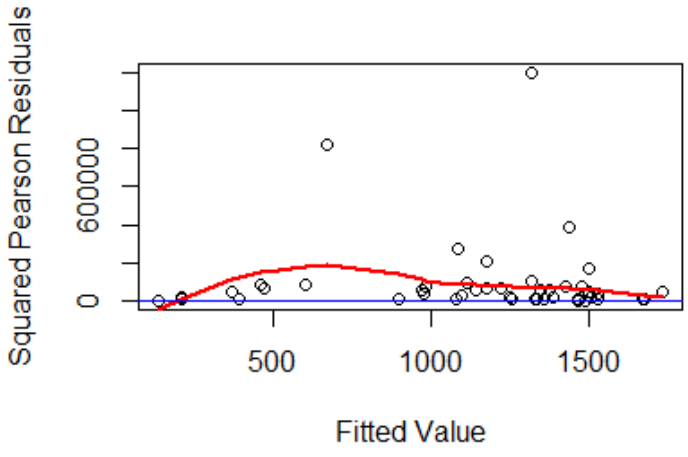

Figure 6: Squared pearson residuals against the fitted value: model 1

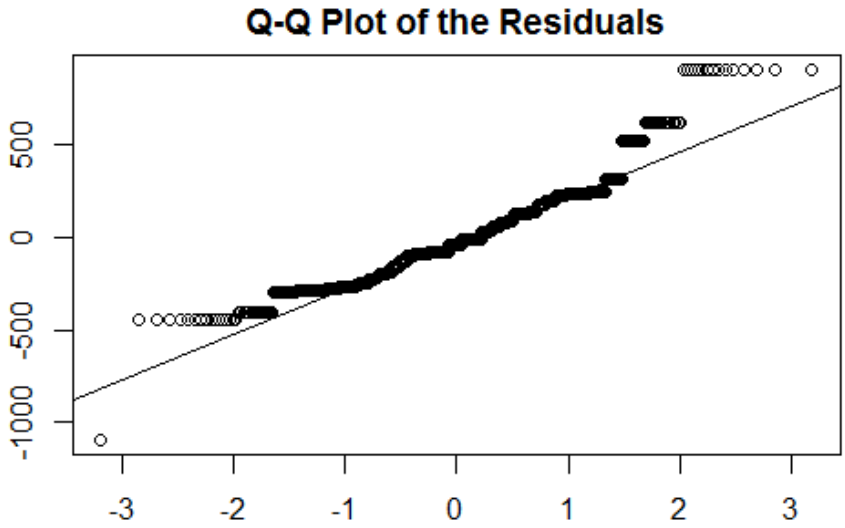

Figure 7: Q-Q plot of residuals for checking normality of residuals: model 1

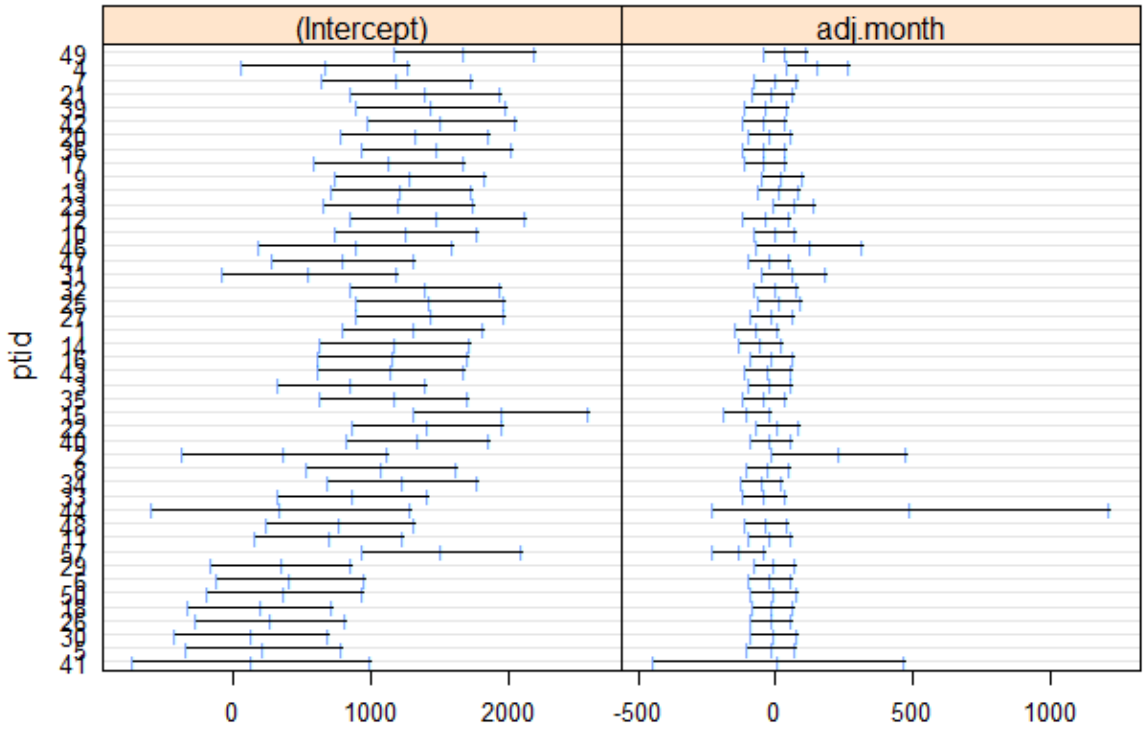

Figure 8: Individual random intercept and random slope for the variation of serum beta-carotene levels

6.2 Tables

	2.5%	25%	50%	75%	97.5%
intcept	-23.208	-5.181	-0.677	4.928	12.868
month	61.562	75.796	83.062	90.439	106.215
vite	165.804	281.898	335.549	393.791	492.343
ind.car	-234.046	-149.417	-102.979	-60.614	20.286
male	-240.430	-145.643	-105.873	-64.597	19.018
bmi	4.050	20.269	27.285	35.510	52.894
cor.int.slope	0.866	0.995	0.999	0.999	1.000
sqrt.D11	106.420	145.810	180.605	223.000	316.252
sqrt.D22	3.251	7.792	13.151	16.776	24.095

Table 6: Results of MCMC fit of linear mixed effect model: secondary analysis

	ptid	month	bcarot	vite	dose	age	male	bmi	chol	cauc	vauc
7	1	6	1228	9.970	30	56	0	24.030	251.000	1100.411	9.079
8	1	6	2088	9.590	30	56	0	24.030	251.000	1100.411	9.079
140	10	6	1465	9.690	45	51	1	23.648	254.500	1414.384	8.495
141	10	6	1674	9.340	45	51	1	23.648	254.500	1414.384	8.495
179	12	12	2507	5.390	45	56	1	23.149	216.000	1365.513	8.658
180	12	12	2401	6.370	45	56	1	23.149	216.000	1365.513	8.658
190	13	6	1074	9.730	45	60	0	21.659	169.000	1475.939	7.258
191	13	6	1507	8.900	45	60	0	21.659	169.000	1475.939	7.258
271	18	7	156	6.420	0	52	1	29.350	225.000	178.108	7.816
272	18	7	148	7.520	0	52	1	29.350	225.000	178.108	7.816
276	18	11	123	7.370	0	52	1	29.350	225.000	178.108	7.816
277	18	11	153	7.090	0	52	1	29.350	225.000	178.108	7.816
325	22	11	2214	6.040	15	62	0	21.657	211.000	1604.824	7.977
326	22	11	2155	4.590	15	62	0	21.657	211.000	1604.824	7.977
342	23	11	2663	5.800	45	60	0	21.673	189.000	1635.509	7.682
343	23	11	2902	5.750	45	60	0	21.673	189.000	1635.509	7.682
359	25	11	2279	9.080	30	58	0	22.361	239.000	1702.715	9.045
360	25	11	2811	7.750	30	58	0	22.361	239.000	1702.715	9.045
403	29	6	403	8.630	0	64	0	19.677	214.000	312.228	7.816
404	29	6	486	7.560	0	64	0	19.677	214.000	312.228	7.816
409	29	11	173	6.920	0	64	0	19.677	214.000	312.228	7.816
410	29	11	211	6.600	0	64	0	19.677	214.000	312.228	7.816
415	30	0	136	5.540	0	60	1	27.103	209.000	125.362	6.057
416	30	0	100	5.200	0	60	1	27.103	209.000	125.362	6.057
428	30	12	73	5.340	0	60	1	27.103	209.000	125.362	6.057
429	30	12	56	5.230	0	60	1	27.103	209.000	125.362	6.057
472	33	11	516	4.030	15	53	1	27.725	171.000	827.359	5.987
473	33	11	544	4.600	15	53	1	27.725	171.000	827.359	5.987
547	40	6	1371	9.190	15	52	0	31.684	228.000	1430.345	7.668
548	40	6	1785	8.510	15	52	0	31.684	228.000	1430.345	7.668
553	40	11	2212	6.050	15	52	0	31.684	228.000	1430.345	7.668
554	40	11	2227	5.290	15	52	0	31.684	228.000	1430.345	7.668
588	43	6	1330	9.650	30	53	0	24.140	212.000	1216.893	8.518
589	43	6	951	9.480	30	53	0	24.140	212.000	1216.893	8.518

Table 5: Table of repeated measurements within a month: note subject with ID 1 has an abnormal measurement at month 6